ToothGrowth Data Analysis Report

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.6.2
```

```
#load data
data("ToothGrowth") # DataSet

head(ToothGrowth)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

```
dim(ToothGrowth)
```

```
## [1] 60  3
```

```
str(ToothGrowth)
```

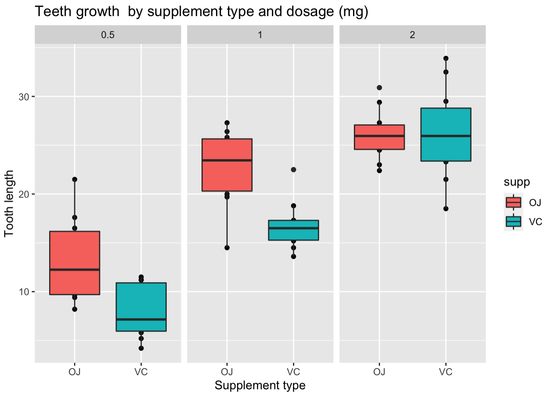
```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

The data has 60 observations with 3 variables: 1) len (numerical)- length of teeth 2) Supplement (OJ or VC) 3) Dose - numeric in milligrams.

```
summary(ToothGrowth)
```

```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

```
qplot(supp,len,data=ToothGrowth, facets=~dose, main="Teeth growth  by supplement type and dosage (mg)",xlab="Supplement type", ylab="Tooth length") + geom_boxplot(aes(fill = supp))
```



There is a positive correlation between the dosage and the tooth growth.

Assumptions The teeth growth follows normal distribution

The variables are independent and identically distributed.

Supplement hypothesis (VC vs OJ) Null hypothesis - there is more teeth growth when using OJ than VC supplement.

A T test will be performed.

Let's obtain teeth growth by supplement from the data:

```
OJ = ToothGrowth$len[ToothGrowth$supp=='OJ']
VC =  ToothGrowth$len[ToothGrowth$supp=='VC']
```

One tailed T test will be performed

```
t.test(OJ, VC, alternative ='greater', paired = FALSE,var.equal = FALSE, conf.level=0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  OJ and VC
## t = 1.9153, df = 55.309, p-value = 0.03032
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.4682687       Inf
## sample estimates:
## mean of x mean of y
##  20.66333  16.96333
```

Since the p value is lower than .005, we reject the null hypothesis.

Dose hypothesis (VC vs OJ)

Null hypothesis- dosage amounts do not make any difference in teeth growth.

```
halfDose = ToothGrowth$len[ToothGrowth$dose == 0.5]
oneDose = ToothGrowth$len[ToothGrowth$dose == 1]
twoDose = ToothGrowth$len[ToothGrowth$dose == 2]
```

We will perform t-test (one tailed).

```
t.test(halfDose, oneDose, alternative='less',paired=FALSE, var.equal=FALSE,
       conf.level=0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  halfDose and oneDose
## t = -6.4766, df = 37.986, p-value = 6.342e-08
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##       -Inf -6.753323
## sample estimates:
## mean of x mean of y
##    10.605    19.735
```

Since p value is lower than 0.05, we reject the null hypothesis.

```
t.test(oneDose, twoDose, alternative = "less", paired = FALSE, var.equal = FALSE, conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  oneDose and twoDose
## t = -4.9005, df = 37.101, p-value = 9.532e-06
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -4.17387
## sample estimates:
## mean of x mean of y
##    19.735    26.100
```

Conclusion

Since p value is also lower than 0.05, the null hypothesis will be rejected. Since its value is very low, it is possible to say that the higher the dosage the longer the tooth will grow.